PCT/AU99/00434

REC'D 13 AUG 1999

WIPO PCT

AU99/434

09/701926

Patent Office
Canberra

5

I, KIM MARSHALL, MANAGER EXAMINATION SUPPORT AND SALES, hereby certify that the annexed is a true copy of the Provisional specification in connection with Application No. PP 6169 for a patent by THE UNIVERSITY OF QUEENSLAND filed on 25 September 1998.

COMMONWEALTH OF AUSTRALIA
AUSTRALIA
PATENT OFFICE

WITNESS my hand this First
day of July 1999

KIM MARSHALL
MANAGER EXAMINATION SUPPORT AND SALES

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

THIS PAGE BLANK (USPTO)

Regulation 3.2

THE UNIVERSITY OF QUEENSLAND

AUSTRALIA

Patents Act 1990

PROVISIONAL SPECIFICATION

for the invention entitled:

"EXPRESSION MODULATING SEQUENCES - III"

The invention is described in the following statement:

EXPRESSION MODULATING SEQUENCES-III

FIELD OF THE INVENTION

The present invention relates generally to novel nucleic acid molecules capable of increasing expression of nucleotide sequences in eukaryotic cells. The novel nucleic acid molecules of the present invention may be used to increase and/or stabilise or otherwise facilitate expression of nucleotide sequences resulting in the presence of a translation product or may be used to down regulate expression by, for example, promoting transcript degradation *via* mechanisms such as co-suppression. The nucleotide sequence of the present invention is referred to herein as an "expression modulating sequence" and generally results in the acquisition of a phenotypic trait or loss of a phenotypic trait. The expression modulating sequence of the present invention is useful *inter alia* to increase and/or stabilise or otherwise facilitate expression of nucleotide sequences in eukaryotic cells and in particular the expression of therapeutically, agriculturally and economically important transgenes. The expression modulating sequence of the present invention may also be used to inhibit, reduce or otherwise down regulate expression of a nucleotide sequence such as a eukaryotic gene including a pathogen gene, the expression of which, results in an undesired phenotype.

BACKGROUND OF THE INVENTION

Recombinant DNA technology is now an integral part of strategies to generate genetically modified eukaryotic cells. For example, genetic engineering has been used to develop varieties of plants with commercially useful traits and to produce mammalian cells which express a therapeutically useful gene or to suppress expression of an unwanted gene. Transposons have played an important part in the genetic engineering of plant cells and some non-plant cells to provide *inter alia* tagged regions of genomes to facilitate the isolation of genes by recombinant DNA techniques.

The maize transposon *Activator (Ac)* and its derivative *Dissociation (Ds)* comprise one of the first transposon systems to be discovered (1,2) and was first used to clone genes by Fedoroff *et al* (3). The behaviour of *Ac* in maize has been studied extensively and excision occurs in both somatic and germline tissue. Studies have highlighted two important features of *Ac/Ds* for tagging. First, the transposition frequency and second, the preference of *Ac/Ds* for transposition in linked sites.

The use of the *Ac/Ds* system has been hampered by the difficulty of data interpretation due, for example, to the high activity of *Ac* in certain plants and insertions at unlinked sites arising from multiple transpositions rather than by a single event from the T-DNA. This problem was addressed by Jones *et al* (4), Carroll *et al* (5) and others where a two component *Ac/Ds* system was developed. In this system, the *Ds* elements were made by replacing the *Ac* transposase gene with a marker gene thereby rendering it non-autonomous. T-DNA regions of binary vectors were constructed by Carroll *et al* (5) and Scofield *et al* (6) carrying either a *Ds* element or a stabilised Activator transposase gene *(sAc)*. The *Ds* element contained a reporter gene (eg. *nos:BAR)* which was shown to be inactivated on crossing with plants carrying the *sAc* (5). This is referred to as transgene silencing. It has been shown that transgene silencing is a more general phenomenon in transgenic plants (7, 8, 9). Many different types of transgene silencing have now been reported in the literature and include: co-suppression of a transgene and a homologous endogenous plant gene (10), inactiviation of ectopically located homologous transgenes in transgenic plants (7), the silencing of transgenes leading to resistance to virus infection (11) and inactivation of transgenes inserted in maize transposons in transgenic tomato (5).

Gene silencing undoubtedly reflects mechanisms of great importance in the understanding of plant gene regulation. It is of particular importance because it represents a severe obstacle to stable and high level expression of economically important transgenes (7).

In work leading up to the present invention, the inventors sought to identify nucleotide sequences which might prevent or otherwise reduce gene silencing and to facilitate increased and/or stabilized gene expression in eukaryotic cells such as plant cells. In accordance with the present invention, the subject inventors have now identified and isolated novel nucleotide sequences

referred to herein as "expression modulating sequences" or "EMSs" which are useful in increasing or stabilizing nucleotide sequence expression in eukaryotic cells such as plant cells. Such increased and stabilised nucleotide sequence expression can also lead to the promotion or induction of transcript degradation *via* mechanisms such as co-suppression. Accordingly, the EMSs of the present invention may also be used to inhibit, reduce or otherwise down-regulate expression of target nucleotide sequences.

SUMMARY OF THE INVENTION

Throughout this specification, unless the context requires otherwise, the word "comprise", or variations such as "comprises" or "comprising", will be understood to imply the inclusion of a stated element or integer or group of elements or integers but not the exclusion of any other element or integer or group of elements or integers.

Sequence Identity Numbers (SEQ ID NOs.) for the nucleotide sequences referred to in the specification are defined following the bibliography. A summary of the SEQ ID NOs is given in Table 1.

One aspect of the present invention provides an isolated nucleic acid molecule comprising a sequence of nucleotides which modulates expression of a second nucleotide sequence inserted proximal to said first mentioned nucleotide sequence.

More particularly, the present invention is directed to an isolated nucleic acid molecule comprising a sequence of nucleotides which increases or enhances expression of a second nucleotide sequence inserted within said first mentioned nucleotide sequence.

Another aspect of the present invention relates to an expression modulating sequence (EMS) comprising a sequence of nucleotides which increases or enhances expression of a nucleotide sequence inserted adjacent to, within or otherwise proximal to said EMS.

Still another aspect of the present invention contemplates a genetic construct comprising an EMS as herein defined and means to facilitate insertion of a nucleotide sequence within, adjacent to or otherwise proximal to said EMS.

Still yet another aspect of the present invention provides a genetic construct comprising an EMS as herein defined and means to facilitate insertion of a nucleotide sequence within, adjacent to or otherwise proximal with said EMS and operably linked to a promoter.

Another aspect of the present invention contemplates a method of increasing or stabilizing expression of a nucleotide sequence or otherwise preventing or reducing silencing of a nucleotide sequence in a eukaryotic cell said method comprising introducing into said eukaryotic cell the nucleotide sequence flanked by, adjacent to or otherwise proximal to an EMS.

More particularly, the present invention provides a method of increasing of stabilizing expression of a nucleotide sequence or otherwise preventing or reducing silencing of a nucleotide sequence in a plant or cells of a plant said method comprising introducing into said plant or plant cells the nucleotide sequence flanked by, adjacent to or otherwise proximal with an EMS.

In an alternative embodiment, the present invention provides a method of inhibiting, reducing or otherwise down-regulating expression of a nucleotide sequence in a eukaryotic cell, said method comprising introducing into said eukaryotic cell the nucleotide sequence flanked by, adjacent to or otherwise proximal with an EMS.

More particularly, the present invention is directed to a method of inhibiting, reducing or otherwise down-regulating expression of a nucleotide sequence in a plant or cells of a plant said method comprising introducing into said plant or plant cells the nucleotide sequence flanked by, adjacent to or otherwise proximal with an EMS.

Yet another aspect of the present invention provides a transgenic animal or plant carrying a nucleotide sequence flanked by, adjacent to or otherwise proximal with an EMS.

Still a further aspect of the present invention provides an improved transposon tagging system, said system comprising a transposable element carrying a nucleotide sequence flanked by, adjacent to or otherwise proximal with an EMS.

TABLE 1

SUMMARY OF SEQ ID NOs.

SEQ ID NO.	DESCRIPTION
1	Nucleotide sequence of tomato α-amylase gene promoter
2	Nucleotide sequence of α-amylase gene promoter
3	Nucleotide sequence of genomic DNA upstream of *Dem* gene followed by *Dem* cDNA coding sequence.
4	Nucleotide sequence upstream of *Ds* insertion (ie. upstream of the *nos:BAR* gene) in a putative patatin gene in tomato
5	Nucleotide sequence downstream of *Ds* insertion (ie. downstream of the *nos:BAR* gene) in a putative patatin gene in tomato
6	Nucleotide sequence of portion of putative tomato homologue of potato patatin gene
7	Nucleotide sequence of portion of potato patatin gene

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 is a diagrammatic representation showing T-DNA regions of binary vectors carrying a *Ds* element (SLJ1561) of the transposable gene (SLJ10512)[5]. The *Ds* element carries a *nos:BAR* gene and is inserted into a *nos:SPEC* excision marker. The transposon gene *sAc* is linked to a 2':*Gus* reporter gene.

Figure 2 is a diagrammatic representation showing an experimental strategy for generating tomato lines carrying transposed *Ds* elements (5). F1 plants heterozygous for both the *Ds* and *sAc* T-DNAs are test-crossed to produce TC_1 progeny. The TC_1 progeny are then screened for lines carrying a transposed *Ds* and a reactivated *nos:BAR* gene.

Figure 3 is a photographic representation showing expression and silencing of the *nos:BAR* gene in various tomato lines. Seedlings were germinated in the presence of phosphinothricin for several weeks and then photographed. A. 1561E, B. UQ406, C. Non-transformed (i.e. does not carry the *nos:BAR* gene), D-F. Three tomato lines that carry silent *nos:BAR* genes.

Figure 4 is a representation showing methylation of a genetically engineered *Ds* transposon in transgenic tomato. Two separate Southern analyses were conducted on 7 individual genotypes; genomic DNA was extracted from leaf tissue (5). The restriction enzymes and probes (shaded boxes) used are shown on the figure. Lanes: 1. Non transformed (i.e. no *Ds* or *nos:BAR* gene), 2. 1561E which carries an active *nos:BAR* gene (due to the fact that it has never been exposed to the transposase gene), 3-6. Four tomato lines that carry silent *nos:BAR* genes, 7. UQ406 which carries an active *nos:BAR* gene due to insertion of the *Ds* in the α-amylase promoter. The enzymes *Sst*II (abbreviated Ss) and *Not*I (abbreviated Nt) are methylation sensitive, whereas *Bst*YI (abbreviated Bs) and *Eco*RI (abbreviated RI) are not. The expected size fragment for unmethylated DNA is indicated by the arrow; larger fragments (as in the silent lines) indicate methylation of the DNA at the *Sst*II or *Not*I sites.

Figure 5 is a representation showing a sequence comparison between the potato α-amylase promoter (15) [SEQ ID NO:2] and the tomato α-amylase promoter [SEQ ID NO:1]. The location of the UQ406 insertion is shown.

Figure 6 is a representation of a nucleotide sequence [SEQ ID NO:3] of genomic DNA from 651 bp upstream of the *Ds* insertion in UQ406 to the beginning of the *Dem* coding sequence, followed by the *Dem* cDNA sequence from the ATG start site at base pair 4097. The target sequences of the *Ds* insertion in UQ406 and *Dem* ATG are underlined. The *Dem* cDNA sequence is shown in italics and underlined.

Figure 7 is a photographic representation showing a stable mutant and a somatic revertant of the *Dem* locus. The seedling at the right in the background is homozygous for the *Ds* insertion in the *Dem* gene. The stable mutant fails to develop beyond the stage shown in the figure. The somatic revertant in the foreground is homozygous for the *Ds* insertion at the zygotic stage of development, but it also inherited a transposase gene which causes *Ds* excision and reversion of the phenotype to wild-type. Somatic revertants are characterized by abnormal cotyledons but develop a functional shoot meristem due to *Ds* excision and restoration of *Dem* function. Each somatic revertant represents an independent transposition event.

Figure 8 is a diagrammatic representation showing an improved transposon tagging strategy using *Dem* as excision marker. The *sAc* and *Ds* parent lines are represented by the upper left and right boxes, respectively. Because the *sAc* is linked to the *dem* mutant +7 allele, somatic revertants can theoretically occur at about the frequency of 1 out of 4 in the F1 progeny. Each somatic revertant represents an independent transposition event. Chr4, chromosome 4 of tomato.

Figure 9 is a diagrammatic representation showing plant expression vector pZorz carrying *Osa:Luc* (12).

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

The present invention is predicated in part on the elucidation of the molecular basis of transposase-mediated silencing of genetic material located within a transposable element. Although, in accordance with the present invention, the molecular basis of gene silencing has been determined with respect to plant selectable marker genes within the *Ds* element of the *Ds/Ac* maize transposon system, the present invention clearly extends to the silencing of any nucleotide sequence and in particular a transgene and to mechanisms for alleviating gene silencing. In accordance with the present invention, nucleotide sequences have been identified which alleviate gene silencing and which increase or stabilise expression of genetic material. Furthermore although the present invention is particularly exemplified in relation to plants, it extends to all eukaryotic cells such as cells from mammals, insects, yeasts, reptiles and birds.

Accordingly, one aspect of the present invention provides an isolated nucleic acid molecule comprising a sequence of nucleotides which increases or stabilizes expression of a second nucleotide sequence inserted proximal to said first mentioned nucleotide sequence.

The term "proximal" is used in its most general sense to include the position of the second nucleotide sequence near, close to or in the genetic vicinity of the first mentioned nucleotide sequence. More particularly, the term "proximal" is taken herein to mean that the second nucleotide sequence precedes, follows or is flanked by the first mentioned nucleotide sequence. Preferably, the second nucleotide sequence is within the first mentioned nucleotide sequence and, hence, is flanked by portions of the first nucleotide sequence. Generally, the second nucleotide sequence is flanked by up to about 10 kb either side of first mentioned nucleotide sequence, more preferably up to about 5 kb, even more preferably up to about 4 kb either side of said first mentioned nucleotide sequence and even more preferably up to about 10 bp to about 1 kb.

Accordingly, another aspect of the present invention is directed to an isolated nucleic acid molecule comprising a sequence of nucleotides which stabilises, increases or enhances expression of a second nucleotide sequence inserted into, flanked by, adjacent to or otherwise proximal to the said first mentioned nucleotide sequence.

The term "expression" is conveniently determined in terms of desired phenotype. Accordingly, the expression of a nucleotide sequence may be determined by a measurable phenotypic change involving transcription and translation into a proteinaceous product which in turn has a phenotypic effect or at least contributes to a phenotypic effect. Alternatively, expression may involve induction or promotion of transcript degradation such as during co-suppression resulting in inhibition, reduction or otherwise down-regulation of translatable product of a gene. In the latter case, the nucleic acid molecules of the present invention may result in production of sufficient transcript to induce or promote transcript degradation. This is particularly useful if a target endogenous gene is to be silenced or if the target sequence is from a pathogen such as a virus, bacterium, fungus or protozoan. In all instances "expression" is modulated but the result is conveniently measured as a phenotypic change resulting from increased or stabilised production of transcript, resulting in increased or stabilised translation product or increased or enhanced transcript production leading to transcript degradation such as in co-suppression resulting in loss of translation product.

The second mentioned nucleotide sequence is preferably an exogenous nucleotide sequence meaning that it is either not normally indigenous to the genome of the recipient cell or has been isolated from a cell's genome and then re-introduced into cells of the same plant or animal, same species of plant or animal or a different plant or animal. More preferably, the exogenous sequence is a transgene or a derivative thereof which includes parts, portions, fragments and homologues of the gene.

The first mentioned nucleotide sequence described above is referred to herein as an "expression modulating sequence" (EMS) since it functions to and is capable of increasing or stabilizing expression of an exogenous nucleotide sequence such as a transgene or its derivatives. This in turn may have the effect of alleviating silencing of an exogenous nucleotide sequence or may promote transcript degradation such as *via* co-suppression. The latter is particularly useful as a defence mechanism against pathogens such as but not limited to plant viruses and animal pathogens.

Accordingly, another aspect of the present invention relates to an expression modulating sequence (EMS) comprising a sequence of nucleotides which increases, enhances or stabilizes expression of a second nucleotide sequence inserted within, adjacent to or otherwise proximal to said EMS.

The term "modulating" is used to emphasise that although transcription may be increased or stabilised, this may have the effect of either permitting stabilised or enhanced translation of a product or inducing transcription degradation such as *via* co-suppression.

The EMSs of the present invention were identified, in accordance with the present invention, following transposon mutagenesis of plants with the *Ds/Ac* transposon system. The *Ds* element carries a reporter gene *(nos:BAR)* which is normally silenced upon exposure to the transposase gene. In a few cases, plants are detected in which *nos:BAR* expression is not silenced. In accordance with the present invention, it has been determined that the *Ds* element inserts within, adjacent to or otherwise proximal with an EMS which results in increased or stabilized expression of the *nos:BAR*. In other words, the EMS facilitates expression of a gene and preferably an exogenous gene or a transgene. This in turn may result in gene product being produced or induction of transcript degradation such as *via* co-suppression.

The EMSs of the present invention are conveniently provided in a genetic construct.

Accordingly, another aspect of the present invention contemplates a genetic construct comprising an EMS as herein defined and means to facilitate insertion of a nucleotide sequence within, adjacent to or otherwise proximal with said EMS.

The term "genetic construct" is used in its broadest sense to include any recombinant nucleic acid molecule and includes a vector, binary vector, recombinant virus and gene construct.

The means to facilitate insertion of a nucleotide sequence include but are not limited to one or more restriction endonuclease sites, homologous recombination, transposon insertion, random insertion and primer and site-directed insertion mutagenesis. Preferably, however, the means is

one or more restriction endonuclease sites. In the case of the latter, the nucleic acid molecule is cleaved and another nucleotide sequence ligated into the cleaved nucleic acid molecule.

Preferably, the inserted nucleotide sequence is operably linked to a promoter in the genetic construct.

According to this embodiment, there is provided a genetic construct comprising an EMS as herein defined and means to facilitate insertion of a nucleotide sequence within, adjacent to or otherwise proximal with said EMS and operably linked to a promoter.

Conveniently, the genetic construct may be a transposable element such as but not limited to a modified form of *Ds*. A modified form of *Ds* includes a *Ds* molecule comprising an EMS and a nucleotide sequence such as but not limited to a reporter gene, a gene conferring a particular trait on a plant cell or a plant regenerated from said cell or a gene which will promote co-suppression of an endogenous gene.

Another aspect of the present invention contemplates a method of increasing or stabilising expression of a nucleotide sequence or otherwise preventing or reducing silencing of a nucleotide sequence or promoting transcription degradation of an endogenous gene in a plant or animal or cells of a plant or animal, said method comprising introducing into said plant or animal or plant or animal cells said nucleotide sequence flanked by, adjacent to or otherwise proximal with an EMS.

In an alternative embodiment, there is provided a method of inhibiting, reducing or otherwise down-regulating expression of a nucleotide sequence in a plant or animal or cells of a plant or animal, said method comprising introducing into said plant or animal or plant or animal cells the nucleotide sequence flanked by, adjacent to or otherwise proximal with an EMS.

Yet another aspect of the present invention provides a transgenic plant or animal carrying a nucleotide sequence flanked by, adjacent to or otherwise proximal to an EMS. As a consequence of the EMS, the expression of the exogenous nucleotide sequence is increased or

stabilised resulting in expression of a phenotype or loss of a phenotype.

Although not intending to limit the present invention to any one theory or mode of action, the EMS is proposed to comprise a methylation resistance sequence. A methylation resistance sequence is one which may de-methylate and/or prevent or reduce methylation of a nucleotide sequence such as a target nucleotide sequence.

According to this aspect, the DNA methylation resistant sequence may prevent inhibition of transcription or delay mRNA transcript turnover. This can enhance, increase or stabilise an transcript and translation into a gene product or may induce or promote transcript degradation such as *via* co-suppression.

The present invention further provides for an improved transposon tagging system.

One system employs a modified *Ds* element which now carries an EMS.

Accordingly, another aspect of the present invention is directed to an improved transposon tagging system, said system comprising a transposable element carrying a nucleotide sequence flanked by, adjacent to or otherwise proximal with an EMS.

Another new system employs the *Dem* gene or its derivatives as an excision marker. Reference to "derivatives" include reference to mutants, parts, fragments and homologues of *Dem* including functional equivalents. The *Dem* gene is required for cotyledon development and shoot and root meristem function. Stable *Ds* insertion mutants of *Dem* germinate but fail to develop any further. However, unstable mutants in the *Dem* locus result in excision of the *Ds* element and reversion of the *Dem* locus to wild-type, thereby restoring function to the shoot meristem. In accordance with the present invention, the new system enables selection for transposition.

In accordance with the improved method, transposition is initiated by crossing a *Ds* line with a stabilized *Ac (sAc)* line. The *Ds* line is heterozygous for a *Ds* insertion in the *Dem* gene and the *sAc* line is heterozygous for a stable mutation in the *Dem* gene. A particularly useful mutant in

the *Dem* gene is a frameshift mutation. Both of the *Ds* and *sAc* containing plant lines are wild-type due to the recessive nature of the *Ds* insertion and mutant alleles. The F_1 progeny derived from crossing the *Ds* and *sAc* lines segregate at a ratio of 3 wild-types to 1 mutant. Because the *sAc* is linked to the frameshift *dem* allele, almost all of the F_1 mutants also inherit the transposase gene and can undergo somatic reversion. These revertant individuals have abnormal cotyledons, but *Ds* excision from the *Dem* gene restores function to the shoot apical meristem. Each somatic revertant represents an independent transposition event from the *Dem* locus. By screening for expression of a gene resident on the *Ds* element (e.g. *nos:BAR),* the identification of EMSs is readily determined.

The present invention also provides *in vivo* bioassays for expressed transgenes. The bioassays identify nucleotide sequences which prevent transgene silencing.

In one aspect, the plant expression vector pZorz (see Figure 5) carries a firefly luciferase reporter gene *(luc),* under the control of the *Osa* promoter (12). After bombardment, the gene is expressed in embryogenic sugarcane callus. However, it becomes completely silenced upon plant regeneration. The silencing appears to be correlated with methylation of the transgene. Genetic sequences flanking reactivated *nos:BAR* insertions are inserted in the pZorz vector at the *Hin*dIII site upstream from the *Osa* promoter. These modified pZorz constructs are then used with a transformation marker to transform sugarcane in order to test whether the plant sequences are capable of alleviating silencing of the *luc* gene upon plant regeneration. Restriction endonuclease fragments capable of alleviating silencing of the *luc* gene are subcloned by deletion analysis into smaller fragments to define the sequence more accurately.

In another aspect, a plant expression vector is constructed for testing the EMSs in *Agrobacterium*-transformed *Arabidopsis.* EMSs are placed upstream of the *nos:luc* or *nos:gus* gene linked to a transformation marker and used to test whether EMSs stabilise expression of the *nos:luc* or *nos:gus* gene in *Arabidopsis.*

These aspects of the present invention are clearly extendable to assays using other plants and the present invention contemplates the subject assay and plant expression vector for use in a range of plants in addition to sugar cane.

The present invention further described by the following non-limiting Examples.

EXAMPLE 1

Ds/sAc Transposon system

The inventors have previously developed a two component *Ds/sAc* transposon system in transgenic tomato for tagging and cloning important genes from plants (5, 13). The components of the system are shown in Figure 1 and comprise: i) a non-autonomous genetically-engineered *Ds* element (e.g. SLJ1561), and ii) an unlinked transposase gene *sAc* (SLJ10512), required for transposition of the *Ds* element. To activate transposition, the two components are combined by crossing transformants for each component. A plant selectable marker gene, e.g. *nos:BAR,* is inserted into the *Ds* element to enable selection for reinsertion of the elements following excision from the T-DNA (Figure 1). Surprisingly, the marker gene is irreversibly inactivated when the *Ds* line is crossed to a transformant expressing the transposase gene (5). Silencing occurred when the *Ds* element remained in the T-DNA, and also occurred in the great majority of cases when the *Ds* element transposed to a new location in the tomato genome. None of the other marker genes in the T-DNA is silenced. The silenced marker gene has been shown to be stably inherited, even after the transposase gene segregates away from the *Ds* element in subsequent generations.

EXAMPLE 2

Transposon tagging of a chromosomal region enabling full expression of the *nos:BAR* transgene

The experimental strategy for generating tomato lines carrying transposed *Ds* elements from T-DNA 1561E is shown in Figure 2. The *Ds* element in 1561E carries a *nos:BAR* marker gene. In construction of the *Ds,* the 5' end of the *nos* promoter is cloned into the *Xho* I site, 1100 bp from the 3′ end of *Ac*. As a strategy to tag regions of the tomato genome associated with high level gene expression, hundreds of plants carrying transposed *Ds* elements are screened for resistance to phosphinothricin (PPT), the selection agent for the *BAR* gene. Several lines are identified which show at least some level of resistance. One line, called UQ406, carries a single transposed *Ds* element (without the transposase gene which has segregated away) and is resistant to PPT (Figure 3). Stable inheritance of *BAR* gene expression in this line has been demonstrated through several generations. These results indicate that the strategy for tagging active chromosomal regions by screening for PPT resistance is a successful approach. Southern hybridization analysis of the original *Ds* transformant 1561E, UQ406 and several lines carrying silenced *nos:BAR* transgenes indicates that silencing is correlated with methylation of the *Sst*II site in the *nos* promoter (Figure 4). Total leaf tissue is used in this analysis, and the *Sst*II site in the *nos* promoter in UQ406 is partially methylated. In silent *nos:BAR* genes, a *Not*I site immediately downstream from the coding sequence is also methylated (Figure 4). In UQ406, the *Not*I site is unmethylated, as in 1561E (Figure 4).

EXAMPLE 3

Cloning sequences flanking active *nos:BAR* genes

GenomeWalker (14) is used to clone the tomato DNA sequences flanking the *Ds* element in UQ406. The DNA flanking the *Ds* element in line UQ406 is cloned and sequenced, and a search of the PROSITE database reveals that the *Ds* has inserted into the promoter region of an α-amylase gene. The promoter [SEQ ID NO:1] shows strong homology to an α-amylase promoter of potato (15; Figure 5) [SEQ ID NO:2] and the coding sequence of the gene has strong homology with one of 3 reported potato α-amylase cDNAs (16). The DNA from 651 bp

upstream of the UQ406 insertion to the end of the *Dem* coding sequence, has been sequenced (Figure 6) [SEQ ID NO:3].

EXAMPLE 4

An improved transposon tagging strategy for transgenic tomato

The inventors have used the transposon tagging system described in Example 1 (also see Figure 2) to tag and clone two important genes involved in shoot morphogenesis. The *DCL* gene is required for chloroplast development and palisade cell morphogenesis (13) and the *Dem* (*D*efective *E*mbryo *M*er stem) gene is required for cotyledon development and shoot and root meristem function. Stable *Ds* insertion mutants of *Dem* germinate but fail to develop any further (Figure 7). Figure 7 also shows an example of an unstable mutant of the *Dem* locus. Upon germination, these variegated seedlings appear at first to be mutant. However, the transposase gene activates transposition of the *Ds* and reversion of the *Dem* locus to wild-type, thereby restoring function to the shoot meristem.

While the transposon tagging system described in Figure 2 has been successful in tagging genes and a chromosomal region alleviating transgene silencing, it does have two associated inefficiencies. First, transposition cannot be selected in the shoot meristem of F_1 plants heterozygous for *Ds* and *sAc*. As a consequence, many TC_1 progeny derived from test-crossing these F_1 plants still have the *Ds* located in the T-DNA. The other limitation of the system is that sibling TC_1 progeny derived from a single F_1 plant often carry the same clonal transposition and reinsertion event. The extent of clonal events amongst sibling TC_1 progeny can only be monitored by time consuming and expensive Southern hybridisation analysis.

These two inefficiencies in the transposon tagging strategy are overcome in accordance with the present invention by using the *Dem* gene as an excision marker. The new system enables selection for transposition in the shoot apical meristem and visual identification of plants carrying independent transposition events. Transposition is initiated by crossing a *Ds* line with a *sAc* line (Figure 8). The *Ds* line is heterozygous for a *Ds* insertion in the *Dem* gene and the *sAc* line is heterozygous for a stable frameshift mutation in the *Dem* gene (Figure 8). The frameshift allele

is derived from a *Ds* excision event from the *Dem* locus. Both the *Ds* and *sAc* lines are wild-type due to the recessive nature of the *Ds* insertion and frameshift alleles. PCR tests on intact leaf tissue have been developed for the rapid identification of these *Ds* and *sAc* parental lines. The F_1 progeny derived from crossing the *Ds* and *sAc* lines segregate at the expected ratio of 3 wild-types to 1 mutant. Because the *sAc* is linked to the frameshift *dem* allele, almost all of the F_1 mutants also inherit the transposase gene *(sAc)* and can undergo somatic reversion. These revertant individuals have abnormal cotyledons, but *Ds* excision from the *Dem* gene restores function to the shoot apical meristem (see Figure 7). Each somatic revertant represents an independent transposition event from the *Dem* locus. A non-destructive test for *nos:BAR* expression is used involving application of PPT (the selective agent for expression of *BAR* gene) to a small area of a leaf. Somatic revertants resistant to PPT are grown though to seed and the F_2 progeny are screened again for PPT resistance. Lines carrying transposed *Ds* elements expressing *nos:BAR* are selected for more detailed molecular analysis. Three independent insertions (UQ11, UQ12 and UQ14) carry active *nos:BAR* genes. The donor *Ds* was originally located in the *Dem* gene (Figure 4) and in that location in the *Dem* gene the *nos:BAR* gene was silent.

The efficient saturation mutagenesis of this chromosomal region is dependent on the use of the *Dem* gene as a selectable marker for independent transposition events. A recombinant selectable marker for independent transpositions is produced and transformed into tomato for saturation mutagenesis in other chromosomal regions of tomato. This system may be introduced into any species possessing the *dem* mutation, in order to facilitate transposon tagging of genes.

EXAMPLE 5

Ds transposon tagging of a putative patatin gene

DNA sequences flanking the active *nos:BAR* in a line designated UQ12 have similarly been cloned and sequenced. The flanking DNA appears to correspond to an intron in a homologous potato patatin gene. Patatin is the major protein in the potato tuber and has many potentially-important characteristics.

The sequence upstream of the *Ds* insertion (i.e. upstream of the *nos:BAR* gene) is as follows:

```
AATCAAAGAG GAATTNAATT CCNCAAAATT TCATCCATAG ATTTTGNGTC     50
TCTGAAAATT AAAGTGACTT TGTAATCTGA AACCTAGAGT CCTCAACCAT    100
ATCATTGACC ATTAAGCCAT ACCCTTAAAT GTAGGGAATT TGAAGTTTTA    150
AAAACCACAC TTTGTTATTT ATTGGCCCAA ATACTCGATA ATCTTTACAT    200
TATTGAAAAT CAACATTCAA AAGGAACGAA CCTTCAATCA CACCATCAAT    250
GTCAACTTTC TTTTATTTTG GATAATCTAA GTTTTTAAAT TGCAGTAAAA    300
TNAAATAAAA CCCTAAACTT CTTCTAGGTT GAGACTTAGT AAATATGAAT    350
TATATAAAGA ATTCATGACA AATGAGACAT AAGAATAGTG CCAGCAAATT    400
ACTTTTTTGA TATCTTATCT GTGATATCGG AATTTTAACT ACCATAAATT    450
TATGAATGAA ATATCACTTA TCTATTAGAG AGGATTTAAT CTCCCTTATA    500
ATGACATTGA TAAAAGCAAG NACAAGTGCT CTTTATTTCT TAATTACAAA    550
TCCTTAAATA GATAAAAGCT ACGAATAACA TAATATCCTT AAATAGATAA    600
AAGCTACGAA TAACATAATA GTATATTACT CCNAATTATT TTGATTTATT    650
TAAAATGACT CCACTAATCC TGATGTGGTC TAGG [SEQ ID NO:4]       684
```

The tomato sequence immediately downstream of the *Ds* insertion (i.e. downstream of the *nos:BAR* gene) is as follows:

```
GGTCTAGGCC CTGGGTCTAG GAAACAAAAT AACTTATTTG ACTCCTAAAC     50
AATAGCAACA TACAAACCAC TGATATTGTA CAAGTAAAAT TCAATAAAAT    100
TCTAGCTCTC TCAAACACTT TTAAAATTGT TATTTCTGTT TTGTCTGTGT    150
CATATTATGA CCTACACAAC AACAACAACA ACGAATTTAG TGAAACTCTA    200
CAAAGTGGAG CCTGAAGTCG AGAGTTTACG CGGGCCTTAT CACTATCTTT    250
TCGAGATAAA AAAATTATTT TTAAAAGATC ATCGACTTAA ACAAACCAAA    300
CAATAATTAA AAAAATATGA ATTAATAGCA AAGCAGTGTG GACCATATAT    350
ACAAAAATCT ATAACAACAA CAAGGTGCAG AGCATTATTC CAACTAAGAT    400
CGAAGTTGTG ATACTGTCAT AATAAAAATG ACACATATTT TGACAACATA    450
AAAAATAAAT AACCATAAAA TATATCATAG AAAAATGAAT ATATTAGAAC    500
AGCTCACTCC AATATTAAAA GAGAGAAAAA AAATATTTTC CCACCACAAT    550
GCCATAATCC TTGAGCTTAG CTATTTATAA GTAAAAAAAA TGTTTTCTTG    600
GATAAATAGA AAAAGAAATA ATAATTAAAC ATAACCAATC ACTTCACAAA    650
TAAGAGTGTA TT   [SEQ ID NO:5]                              662
```

The level of homology between the potato and our tomato sequence is as follows:

```
Tomato:   307 ATTTATTTTTAGGAAAAATTATCTAAATACACATCTTATTTACCATATACTCTAAAAT 248
              | |||| |||||||||||||  |||||||||| |||| | ||  |||| |||||||| |
Potato:  1914 AATTATATTTAGGAAAAATTACATAAATACACAACTTAATATATTATATTCTCTAAAATT 1973

          247 TCC 245    [SEQ ID NO:6]
              |||
         1974 TCC 1976   [SEQ ID NO:7]
```

EXAMPLE 6

Tagging of additional genes

Selecting for transposition of a methylated *Ds* from the *Dem* locus and for expression of the *nos:BAR* gene (i.e.: demethylation of the *Ds*) efficiently identifies *Ds* insertions into genes, as opposed to so-called "junk DNA". The sequences adjacent to five of these *Ds* insertions have been cloned and sequenced, and in all cases the *Ds* insertion is in the vicinity of a known gene.

The five lines carrying active *nos:BAR* genes associated with genes are:

- *Ds* insertion in UQ406 - associated with the promoter of an α-amylase gene (Example 3, above);
- *Ds* insertion in UQ12 - associated with a putative palatin gene (Example 5, above);
- *Ds* insertion in UQ11 - associated with the Right Border of the *Agrobacterium* T-DNA in 1516E (refer to Figure 2). This was the T-DNA carrying the *Ds* that was initially transformed into tomato. In other words, the *Ds* transposed from the *Dem* locus back into the T-DNA;
- *Ds* insertion in UQ14 - associated with or closely linked to a putative sucrose synthase gene; and
- *Ds* insertion in UQ13 - associated with or closely linked to a putative UDP-glucose-pyrophosphorylase gene.

In four of these instances, the *Ds* has inserted into or near sequences homologous to carbon metabolism genes. These data indicate that many C metabolism genes and many so called house-keeping genes contain de-methylation sequences or sequences which prevent or reduce methylation.

EXAMPLE 7

A rapid bioassay for identification of tomato DNA sequences capable of alleviating transgene silencing in a heterologous plant species

An efficient transformation system has been developed for sugarcane, based on particle bombardment of embryogenic alleles, followed by plant regeneration (17). The bioassay is useful for identifying tomato sequences which prevent transgene silencing and employs the plant expression vector pZorz (Figure 9). This plasmid carries a firefly luciferase reporter gene *(luc),* under the control of the *Osa* promoter (12). After bombardment of embryogenic callus of sugar cane, the luciferase gene is expressed as observed by visualisation of the chemiluminescence of the luciferase enzyme. However, it becomes completely silenced upon plant regeneration in normal sugar cane. This is used to test the system. The silencing appears to be correlated with methylation of the transgene. Tomato sequences flanking reactivated *nos:BAR* insertions are inserted in the pZorz vector at the *Hind*III site upstream from the *Osa* promoter (Figure 10). These modified pZorz constructs are then used with a transformation marker to transform sugarcane in order to test whether the tomato sequences are capable of alleviating silencing of the *luc* gene. They are then subcloned by deletion analysis into smaller fragments to more accurately define the sequences.

Tomato sequences flanking reactivated *nos:BAR* insertions are also introduced next to a *nos:BAR, nos:LUC* or *nos:GUS* recombinant gene in another plasmid vector. These modified recombinant *BAR, LUC* and *GUS* genes are inserted into binary vectors (4) for transformation into *Arabidopsis thaliana* (18) to test the ability to prevent silencing of the *nos:BAR* gene in *Arabidopsis.*

EXAMPLE 8

Analysis of sequences responsible for reactivating *nos:BAR* expression

The borders of DNA elements that prevent transgene silencing are initially defined by deletion analysis of clones that yield positive results in the bioassays. The smallest active clone for each chromosomal region is then sequenced and characterised in detail. Sequences from independent *Ds* insertions are compared for homologous DNA elements.

Those skilled in the art will appreciate that the invention described herein is susceptible to variations and modifications other than those specifically described. It is to be understood that the invention includes all such variations and modifications. The invention also includes all of the steps, features, compositions and compounds referred to or indicated in this specification, individually or collectively, and any and all combinations of any two or more of said steps or features.

BIBLIOGRAPHY

1 McClintock, B. (1947) *Carnegie Inst. Washington Year Book 46:* 146-152.

2 McClintock, B. (1948) *Carnegie Inst. Washington Year Book 47:* 155-169.

3 Fedoroff, N. *et al,* (1984) *Proc. Natl. Acad. Sci. USA 81:* 3825-3829.

4 Jones, J. *et al.* (1992) *Transgenic Res. 1:* 285-297.

5 Carroll, B. J. *et al,* (1995) *Genetics 139:* 407-420.

6 Scofield, S. *et al.* (1992) *Plant Cell 4:* 573-582.

7 Finnegan, J and McElroy, D (1994) *Biotech 12:* 883-888.

8 Spiker, S and Thompson, W F (1996) *Plant Physiol 110:* 15-21.

9 Matzke, M A and Matzke, A J M *Plant Physiol 107:* 679-685.

10 Jorgensen, R A (1995) *Science 268:* 686-691.

11 Smith, H A *et al* (1994) *Plant Cell 6:* 1441-1453.

12 Bower, R *et al* (1996) *Mol. Breeding 2:* 239-249.

13 Keddie, J S *et al* (1996) *EMBO J 15:* 4208-4217.

14 Siebert, P.D. *et al.* (1995) *Nucleic Acids Res. 23:* 1087-1088.

15 International Patent Publication No. WO 96/12813.

16 International Patent Publication No. WO 9012876-A.

17 Bower, R and Birch, R G (1992) *Plant J. 2:* 409-416.

18 Bechtold, N. and Bouchez, D. (1995) In: I. Potrykus and G. Spangenberg (eds). Gene transfer to Plants. pp.19-23.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: THE UNIVERSITY OF QUEENSLAND

(ii) TITLE OF INVENTION: EXPRESSION MODULATING SEQUENCES-III

(iii) NUMBER OF SEQUENCES: 7

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: DAVIES COLLISON CAVE
(B) STREET: 1 LITTLE COLLINS STREET
(C) CITY: MELBOURNE
(D) STATE: VICTORIA
(E) COUNTRY: AUSTRALIA
(F) ZIP: 3000

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: AUSTRALIAN PROVISIONAL
(B) FILING DATE: 25-SEP-1998

(vii) PRIOR APPLICATION DATA
(A) APPLICATION NO. PP3901
(B) FILING DATE: 4-JUNE-1998

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: HUGHES, DR E JOHN L
(C) REFERENCE/DOCKET NUMBER: EJH/EK

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: +61 3 9254 2777
(B) TELEFAX: +61 3 9254 2770
(C) TELEX: AA 31787

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1217 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

TTTGAAATTT ATGTATTTAT CTATAGCATT AGAAACTATA AGAGTTGTTA GCTTCACTTG 60

GCTTACTGTT GTGCTCAAAG CAACTTCATC ATCATACAGT ATGGTTTTGA TATGCTCTTC 120

CATTATCACT GAGCCTTATG ATTATGTTTT ACGAGCTTAT AATATCACTG ATGGTGATTC 180

AGTATTGTGA TTATGTCCTT CGTTGATTAT TCTGTTTCAT ACAAGTCGTG TAATTTGCTG 240

TTTGTGACAG TACGATAGAT CGACTCAACC TTCTGAGGTA TTAGTTGAAG TTCATGTAAA 300

TTAGCTTTGT TTATCATAGT AGCATTTGAT TATTGATGCT CTGTAGCTAA TGATAAGCCA 360

TTGGAGGGAA GCAAGCTTTC TAAATGAATC TACGAATGGA TGATAAAGTT CATGAATATT 420

TTTGTTACTT CTGCAGTCAG ATCATGAGTT ATTGAGTCTA TTGTTTTTTT AAGCCTGTTT 480

CAGATGATCC ATCATCAGTA ACAACATACA CGGTGTAGTC CCAAATCCAT CATATGCACC 540

TTCTTTTCTT CAATTTGGTC TTGTTTTTTT TTTTTCATGA TGTCATTGAA TTATTCAAGA 600

AGTCACTTCG AGCATAATGA TTTTTCAAAA TCCACCTTTG TTCAAGCACT ACCACGTCTT 660

TTCATCTAGC CCACAACCGT GGTGGAGGAT CTAGAATTTT CATGAAAGGA TTCAAAATTT 720

ACAAACATAT ATATACACTA TACACTATGA ATCCACTAAT ACTAGATGGT GCACCTGTGC 780

CCCCACTCAT GTGAAAGCCT ATTCTCAATT TTTTATTTTC CACAACTTAA ATACAGACCG 840

CACAACTCCC GTGTCTTGTG TGCTCGTCGC TCAGCATGCA AGTCGAGAAA AGAAAGACCA 900

AAACAATGAA AACTTTACGA AAAATCAAAA AGTTGAAGGA CTTTAACGTC GAGATCTCTC 960

GTAGAAAACC TCTTTTGTAA GGTTGCATAC AATACTTTTT TTTCAGACTT TACTTATGGT 1020

ATTATACTGA ATATGTTATT GCTGTTATAG TAGTTGAGTG ACGTTTGAGG GAATTTCTAG 1080

TCCGTTAATC TTGTACTCAG TGTGTCTACT TTTCAAAAAA GTCAGTTTTT CAGTCTCTAA 1140

AACACATTTA AATAAGAGTT TCTTTGCCCA TCTTTTGTTC CTCATCCTAG GCTTGGAGTC 1200

AACACAACAC AACAACA 1217

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1114 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
TTTGAAATTT ATGTATATAT CTGTAGCATT AGAAACTATA AGAGTTGTTA GCTTCACTTG      60
TCTTATTGTT GTGCTCAAAG CAACTTCATC ATACAGTATG GTTTTTATAT GCTCTTCCAT     120
TATCACCGAA CCTTATGATT ATGTGTACGA GCTTATAATA TTACTGATGG TGATTCAGTA     180
TTATGATTAT GTCCTCCATT AATTATTCTG TTTCATACAA GTCGTGTAAT TTGCTGTTTG     240
TGATTGTACG ATAAATTGAT TCAACCTTCT GCGGTGTTGG TTGAAGTTCA AGTAAATTAG     300
CTTTATTTAT CATAGTAGCA TTTGATTATT GATGCTCTGT AGCTAATGAT AAGCCATTGA     360
AGGGAAGCAG AAATGGTAAA GCTTTCTAAA ATGAATCTAC GAATGGATGA TAAAGTTAAT     420
GAATATTGTT GATACTTCTG CAATCAGATT ATGAGTTACT GAGTCTACTG TTTTTTAAGC     480
CTGTTTCAGA TGATCGATCA TCAACAACAA CATATTCAGT GTAGTAGACA TGATCGATCA     540
CTTTCTAATT TTCGATTATG CACCCTCTTT TCTCCAATTT GGTCGTCTTC TTTTTTTCAT     600
GATGTCACTG AATTATTCTC TGGTCGTCCC CACCATTCAG GAAGTCACTT CGAGCATAAT     660
GTGAAAACAT CCACATTTTT CAAATCCAGC AGAATTTTCA TCAAACGGGG TTCAACATTT     720
ACTACATGTA TACACTCTGA AGTCTGAATC CACTAATTCT AGATGGTGCA TCTGTGCCCC     780
CACACTTGTG AAAGCTTATT CTCAATTTTT TATTTTCCAA CAACTTGAAT TCAGACCACA     840
CAACTCCCGT GTCTTGTACG GTCAGCATCT GAGTGGAGAA CTCAATTAAG TGACTTTAAC     900
GTCGAGTTCT ATAGTAAACA ACCCCTATAT CTTTTTTCAA GCATGTTAAG ATTGCGAACA     960
CACTGAAATT TCCAGGTCGT TAATCTTGTA CCCAGTGTGT GTACTTTTAA AAAAAAAAGT    1020
CAGTTTTTTA GTCTCTAAAA CACATTTAAA TAGAGTTTAT TTGCCATCTT TTGTTCCTCA    1080
TACTAGACTT CGGAGTCAAC ACAACACAAC AACA                                1114
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 6263 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
CGACGGCCCG GGCTGGTAAA TGCGGAAGCT TGTTACAGAT TTGAAATTTA TGTATTTATC     60
TATAGCATTA GAAACTATAA GAGTTGTTAG CTTCACTTGG CTTACTGTTG TGCTCAAAGC    120
AACTTCATCA TCATACAGTA TGGTTTTGAT ATGCTCTTCC ATTATCACTG AGCCTTATGA    180
TTATGTTTTA CGAGCTTATA ATATCACTGA TGGTGATTCA GTATTGTGAT TATGTCCTTC    240
GTTGATTATT CTGTTTCATA CAAGTCGTGT AATTTGCTGT TTGTGACAGT ACGATAGATC    300
GACTCAACCT TCTGAGGTAT TAGTTGAAGT TCATGTAAAT TAGCTTTGTT TATCATAGTA    360
GCATTTGATT ATTGATGCTC TGTAGCTAAT GATAAGCCAT TGGAGGGAAG CAAGCTTTCT    420
AAATGAATCT ACGAATGGAT GATAAAGTTC ATGAATATTT TTGTTACTTC TGCAGTCAGA    480
TCATGAGTTA TTGAGTCTAT TGTTTTTTTA AGCCTGTTTC AGATGATCCA TCATCAGTAA    540
CAACATACAC GGTGTAGTCC CAAATCCATC ATATGCACCT TCTTTTCTTC AATTTGGTCT    600
TGTTTTTTTT TTTTCATGAT GTCATTGAAT TATTCAAGAA GTCACTTCGA GCATAATGAT    660
TTTTCAAAAT CCACCTTTGT TCAAGCACTA CCACGTCTTT TCATCTAGCC CACAACCGTG    720
GTGGAGGATC TAGAATTTTC ATGAAAGGAT TCAAAATTTA CAAACATATA TATACACTAT    780
ACACTATGAA TCCACTAATA CTAGATGGTG CACCTGTGCC CCCACTCATG TGAAAGCCTA    840
TTCTCAATTT TTTATTTTCC ACAACTTAAA TACAGACCGC ACAACTCCCG TGTCTTGTGT    900
GCTCGTCGCT CAGCATGCAA GTCGAGAAAA GAAAGACCAA AACAATGAAA ACTTTACGAA    960
AAATCAAAAA GTTGAAGGAC TTTAACGTCG AGATCTCTCG TAGAAAACCT CTTTTGTAAG   1020
GTTGCATACA ATACTTTTTT TTCAGACTTT ACTTATGGTA TTATACTGAA TATGTTATTG   1080
CTGTTATAGT AGTTGAGTGA CGTTTGAGGG AATTTCTAGT CCGTTAATCT TGTACTCAGT   1140
GTGTCTACTT TTCAAAAAAG TCAGTTTTTC AGTCTCTAAA ACACATTTAA ATAAGAGTTT   1200
CTTTGCCCAT CTTTTGTTCC TCATCCTAGG CTTGGAGTCA ACACAACACA ACAACAATGA   1260
ATTTCCATTT TTCTGTTTCT TTACTTCTCT CTTTATCTCT TCCTATGTTT GCCTCTTCGA   1320
CGGTGTTATT TCAGGTATCC ATCTCCAAAG AACCTTATTT TTCTCTTAAC TTTTCCTATG   1380
TATATGTATC TCTATGTTTA TGTAGTACTT GCTCAAGTAT ATAAAGAAAA GTTAGTTTCT   1440
CTAGAATCTT TGAATTCATT TGTTAGGGGT TCAATTGGGA TTCGAGTAAT AAGCAAGGCG   1500
GATGGTACAA CTCTCTCATC AACTTAGTTC CGGACTTGGC TAAAGCTGGA GTTACTCATG   1560
TTTGGTTGCC ACCATCATCT CACTCCGTTT CTCCTCAAGG TAATTTTCGG AGTGATTGTG   1620
```

ACCTAGTAAT	CCAATGAAGT	CAAAATAACC	ACGGAAGATT	AGAGTCTAAA	TTTTAATGAA	1680
AATAGTTCAG	ACAAGTTAAT	GACCAACTTA	TATATTAGTT	CAATCCATAA	AATTTGATGT	1740
AGTAGTTACA	AAATGGAATT	GCTTGAAGGC	TTATGCCATG	TTTTATGCCA	GGTTATATGC	1800
CAGGAAGGTT	GTATGACTAG	GATGCTTCCA	AGTTTGGAAA	TCAGCAACAA	CTGAAAACTC	1860
TTATTAAGGC	TTTAACATGA	CCACGGGATC	AAATCGGTTG	CTGATATAGT	GATAAATCAT	1920
AGAACTGCTG	ATAACAAAGA	TAGCAGGGGA	ATATACAGCA	TCTTTGAAGG	AGGAACATCT	1980
GATGACCGGC	TTGATTGGGG	TCCATCTTTC	ATTTGCAGGA	ACGACACACA	ATATTCTGAT	2040
GGCACGGGGA	ATCCAGACAC	GGGTTTGGAC	TTTGAACCTG	CACCTGATAT	CGATCATCTT	2100
AATACGAGAG	TGCAGAAAGA	GTTATCAGAC	TGGATGAACT	GGCTGAAATC	TGAAATTGGA	2160
TTTGATGGTT	GGCGTTTCGA	TTTTGTTAGG	GGATATGCAC	CTTGCATTAC	CAAAATTTAT	2220
ATGGGAAACA	CGTCCCCGGA	TTTTGCTGTT	GGTGAATTGT	GGAACTCTCT	TGCTTATGGC	2280
CAGGACGGGA	AACCGGAATA	TAACCAGGAC	AATCATAGAA	ATGAGCTAGT	TGGTTGGGTA	2340
AAAAATGCGG	GGCGGGCTGT	AACAGCTTTT	GATTTTACAA	CAAAGGGAAT	TCTTCAAGCT	2400
GCAGTTCAAG	AAGAGTTATG	GAGATTGAAG	GATCCCAATG	GAAAACCTCC	TGGGATGATC	2460
GGTGTTTTGC	CTCGAAAAGC	TGTGACTTTT	ATCGATAATC	ATGATACTGG	ATCGACACAA	2520
AATATGTGGC	CTTTCCCTTC	AGACAAAGTT	ATGCAAGGAT	ATGCATACAT	TCTTACTCAT	2580
CCAGGAATCC	CATCCGTGGT	AAAAAAAATA	AATAAATTCT	TTCTACATAT	CTCATTGTTT	2640
TCTATTTTAC	AAGAAATTTA	TATTCTTTTC	CAGGGGATTT	GAGAAACTCG	GCCTGTGGGA	2700
GTTTGCTCAC	ATTGCCAGTC	TCGTAATCCA	TAAACAAACA	CTCAAACTCT	GAGTGTGCAC	2760
ATCTAGACAC	CTCAACTCGT	TTTTCACCGT	GTTAATTGAA	CACTTCAACT	TACAAAATGA	2820
TCGTGTAGCA	CCTCCAAAAA	TTATGTGTCA	CAATTAGCCA	CGTGCGAGAT	ACACGAAAAT	2880
GAGTTGGAGT	AGTTAGTTGC	CAAATAAAAC	CAAGCTGAGG	TGTCTAAATG	TGCACNCTCA	2940
AAGTNGGATG	TTTACTTGGC	AGCTGAGGCC	GAGGCCATGT	TTGANTGTTA	TGCTTATAGG	3000
ATATGACACA	TTTGTTTCCG	ATTAGCTGAG	GANTTGATTA	AATCCTNGTT	TTNGTTNGCA	3060
GTTTNATNAC	CATTNCTTTG	ATNGGGGCTN	CNAGGATGGA	ATTNCAGCAC	TAANCTCTAT	3120
TAGGAAAAGG	AATAGGATTT	GTGCANCAAG	CAATGTGCAA	ATAATGGCTC	CTGATTCTGA	3180
ATCTTTATAT	ANCAATGGAT	CATCACAAAA	TCATTGTCAA	GATTGGACCA	AAACTTGATC	3240
TTGGAAATCT	TATTCCACCT	AATTATGAGG	TGGCAACTTC	TGGACAAGAC	TATGCTGTAT	3300
GGGAGCAAAA	GGCATAATCA	TATTGTACCA	CACTAAAAGG	GACCATGGCC	ACAATGGTTC	3360
TCATTAGTGT	TAATGTTATA	TGATTGAAAA	TGTAATTTAT	ATTGACATAA	TGAAGGCCAA	3420
AAATTCAAGA	AATTATAAAC	AATTCAATAG	TCCTTGCTCA	ATTCACAATT	ACATTATGAC	3480
TTCTCTATTG	CAAACTAGTT	TGGGTCCACA	TTATTGTCTC	CTAAAATTTT	ACAACATTTC	3540
TTAAGGGAAC	TTAATTAGTT	ACAGTGAACA	TATGTTGAAA	TTACCCTTTA	TCCCCTTACA	3600
ATTGATTTAA	TAAATATTTC	CCCTATCCCT	TTGGTAGTTG	GTTAGAGTTA	TAAGTAACGT	3660

AGAGATTAGT TATAAGAGAA TTTATGTATT ATTATGCAGA TGTTTAGTTA TATCGATTTT 3720
AGTTATTTAT ATGTTGATTA TTTCACCTTC AATAATGCAT ATAAAGATGG TAAATGATTG 3780
GATTGATCGA ATTCGAATGA GTTTGAATAT GAACTAATCT TCAAATTTAA TATAAATTTT 3840
TTTTGTCAAC ATCTATAGCC AAACGGCTCC AAAACAATAA ATAATTTACA TTTATTGTAG 3900
TATTTTATTT AAAATGGGAT NTTCCTCATC CCACTTGTAC CAGTTGAAAC CCTAATAATA 3960
AGCCAATCCA ACCGTCAAAA TTACAAATTT TGAAAATTGC GCTCCTCACA GTTCTCCCCT 4020
ATTCAGATTT GATTCATTCT CTTCATTTTT TGTTTTCACA TTTTACCTCT AAATCAACTC 4080
GAGTCCCTTT GTTCAAATGG GTGCTAATCA CAGCCGTGAA GATCTGGAGC TTTCTGATTC 4140
CGAGTCTGAA TCCGAATATG GGTCCGAGTC TCGAACAAGG GAGGAAGAGG AAGACGAAGA 4200
TAACTACTCA GATGCTAAAA CGACGCCGTC TTCCACTGAT CGGAAACAGA GCAAAACCCC 4260
GTCTTCTTTG GATGATGTTG AAGCAAAGCT GAAAGCTTTA AAGCTTAAGT ATGGTACTCC 4320
TCATGCTAAA ACCCCCACAG CGAAAAACGC TGTTAAACTT TACCTTCATG TTGGTGGGAA 4380
CACTGCGAAT TCCAAATGGG TAGTTTCTGA TAAGGTGACA GCTTATTCGT TTGTTAAATC 4440
GGGTAGTGAG GATGGATCGG ATGATGATGA AAATGAAGAA ACTGAGGAGA ATGCTTGGTG 4500
GGTTTTGAAA ATTGGGTCGA AGGTTCGGGC TAAGATTGAT GAGAATTTGC AGCTCAAGGC 4560
ATTTAAGGAG CAGAAAAGGG TGGATTTTGT GGCGAATGGG GTTTGGGCTG TGAGATTCTT 4620
TGGGGAGGAA GAGTATAAGG CGTTCATTGA CTTATATCAG AGCTGTTTGT TTGAGAATAC 4680
TTATGGGTTT GAGGCAAATG ATGAGAATAG AGTTAAGGTG TATGGTAAAG ACTTTATGGG 4740
GTGGGCAAAT CCAGAAGCTG CGGATGATTC AATGTGGGAG GATGCTGGGG ATAGCTTCGC 4800
GAAGAGCCCT GCGTCTGAAA AGAAGACACC TTTGAGGGTT AACCATGATT TGAGGGAGGA 4860
GTTTGAGGAG GCAGCTAAAG GAGGAGCTAT TCAGAGCTTG GCATTAGGTG CGTTGGATAA 4920
TAGTTTTCTT ATAAGTGATT CTGGAATTCA GGTTGTGAGG AACTATACTC ATGGAATAAG 4980
TGGAAAAGGT GTTTGTGTCA ATTTTGATAA GGAAAGGTCT GCTGTACCTA ATTCCACTCC 5040
AAGGAAAGCT CTACTTCTAA GAGCTGAGAC TAATATGCTT CTCATGAGTC CAGTGACTGA 5100
TAGAAAGCCT CACTCTCGGG GATTACATCA GTTTGATATC GAGACTGGGA AGGTTGTTAG 5160
CGAGTGGAAG TTTGAGAAAG ATGGAACTGA TATCACGATG AGGGATATCA CTAATGATAG 5220
CAAAGGAGCT CAGATGGATC CTTCGGGGTC TACTTTCTTA GGGCTAGATG ATAACAGATT 5280
GTGTAGGTGG GATATGCGTG ATCGGCATGG GATGGTCCAG AATCTAGTTG ATGAAAGTAC 5340
TCCTGTGCTG AATTGGACTC AAGGACATCA ATTTTCGAGG GGAACTAACT TTCAGTGCTT 5400
TGCTACTACT GGTGATGGAT CAATTGTTGT TGGTTCACTT GATGGCAAGA TTAGATTGTA 5460
CTCAAGCAGT TCCATGAGAC AGGCTAAAAC TGCTTTTCCA GGCCTTGGTT CTCCTATCAC 5520
TCATGTGGAT GTTACCTATG ATGGGAAGTG GATATTGGGG ACAACTGATA CTTACTTGAT 5580
ATTGATATGC ACCTTGTTTA TCGACAAGAA TGGAACTACT AAGACTGGTT TTGCTGGTCG 5640
CATGGGAAAT AAGATTTCCG CTCCAAGATT GTTAAAGCTA AACCCTCTCG ATTCACATAT 5700

GGCTGGAGCT AACAAGTTCC GCAGTGCTCA ATTTTCATGG GTCACCGAGA ATGGGAAGCA 5760

AGAGCGCCAC CTCGTTGCTA CTGTTGGGAA GTTTAGTGTG ATCTGGAATT TTCAACAGGT 5820

GAAGGATGGT TCTCATGAGT GTTACCAGAA TCAGGTTGGG TTGAAGAGCT GCTATTGTTA 5880

CAAGATAGTC CTAAGAGACG ACTCTATTGT AGAAAGTCGT TTCATGCATG ACAAGTACGC 5940

TGTTTCTGAC TCACCTGAAG CACCACTGGC GGTAGCAACC CCCATGAAAG TCAGCTCATT 6000

CAGCATCTCT AGCAGGCGCT TACAAATTTG AACAATCATT CTGTTCATAT ACGCAACTTA 6060

TTAGATTTAT CTGTAGCAGA ATTAGTGTCT CTCACACTAA GTAGCTTGAA AAACTGCACA 6120

TCTGCAAATC ATTTCCAGTT CAATGTATTA CTACTTTAGT TTAAAAACCT TAAAAGGCAC 6180

TCTTCCAAAT TCTAGGTATC CTCACCTGAC ATTATTATTG TTGTAATAGC TAATTGTTGC 6240

TTGCTCTAAA TCCCCGTTCA ATG 6263

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 684 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

AATCAAAGAG GAATTNAATT CCNCAAAATT TCATCCATAG ATTTTGNGTC 50

TCTGAAAATT AAAGTGACTT TGTAATCTGA AACCTAGAGT CCTCAACCAT 100

ATCATTGACC ATTAAGCCAT ACCCTTAAAT GTAGGGAATT TGAAGTTTTA 150

AAAACCACAC TTTGTTATTT ATTGGCCCAA ATACTCGATA ATCTTTACAT 200

TATTGAAAAT CAACATTCAA AAGGAACGAA CCTTCAATCA CACCATCAAT 250

GTCAACTTTC TTTTATTTTG GATAATCTAA GTTTTTAAAT TGCAGTAAAA 300

TNAAATAAAA CCCTAAACTT CTTCTAGGTT GAGACTTAGT AAATATGAAT 350

TATATAAAGA ATTCATGACA AATGAGACAT AAGAATAGTG CCAGCAAATT 400

ACTTTTTTGA TATCTTATCT GTGATATCGG AATTTTAACT ACCATAAATT 450

TATGAATGAA ATATCACTTA TCTATTAGAG AGGATTTAAT CTCCCTTATA 500

ATGACATTGA TAAAAGCAAG NACAAGTGCT CTTTATTTCT TAATTACAAA 550

TCCTTAAATA GATAAAAGCT ACGAATAACA TAATATCCTT AAATAGATAA 600

AAGCTACGAA TAACATAATA GTATATTACT CCNAATTATT TTGATTTATT 650

TAAAATGACT CCACTAATCC TGATGTGGTC TAGG 684

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 662 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

```
GGTCTAGGCC CTGGGTCTAG GAAACAAAAT AACTTATTTG ACTCCTAAAC     50
AATAGCAACA TACAAACCAC TGATATTGTA CAAGTAAAAT TCAATAAAAT    100
TCTAGCTCTC TCAAACACTT TTAAAATTGT TATTTCTGTT TTGTCTGTGT    150
CATATTATGA CCTACACAAC AACAACAACA ACGAATTTAG TGAAACTCTA    200
CAAAGTGGAG CCTGAAGTCG AGAGTTTACG CGGGCCTTAT CACTATCTTT    250
TCGAGATAAA AAAATTATTT TTAAAAGATC ATCGACTTAA ACAAACCAAA    300
CAATAATTAA AAAAATATGA ATTAATAGCA AAGCAGTGTG GACCATATAT    350
ACAAAAATCT ATAACAACAA CAAGGTGCAG AGCATTATTC CAACTAAGAT    400
CGAAGTTGTG ATACTGTCAT AATAAAAATG ACACATATTT TGACAACATA    450
AAAAATAAAT AACCATAAAA TATATCATAG AAAAATGAAT ATATTAGAAC    500
AGCTCACTCC AATATTAAAA GAGAGAAAAA AAATATTTTC CCACCACAAT    550
GCCATAATCC TTGAGCTTAG CTATTTATAA GTAAAAAAAA TGTTTTCTTG    600
GATAAATAGA AAAAGAAATA ATAATTAAAC ATAACCAATC ACTTCACAAA    650
TAAGAGTGTA TT                                             662
```

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 63 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

```
ATTTATTTTT AGGAAAAATT ATCTAAATAC ACATCTTATT TTACCATATA CTCTAAAAAT    60
TCC                                                                  63
```

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 63 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

AATTATATTT AGGAAAAATT ACATAAATAC ACAACTTAAT ATATTATATT CTCTAAAATT 60

TCC 63

DATED this 25th day of September 1998

THE UNIVERSITY OF QUEENSLAND

By DAVIES COLLISON CAVE

Patent Attorneys for the Applicants

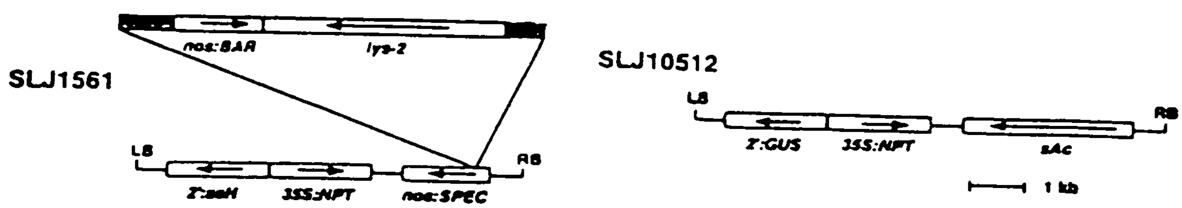
nos:BAR
lys-2
SLJ1561
LB
RB
Z:seH
35S:NPT
nos:SPEC
SLJ10512
LB
RB
Z:GUS
35S:NPT
sAc
1 kb

FIGURE 1

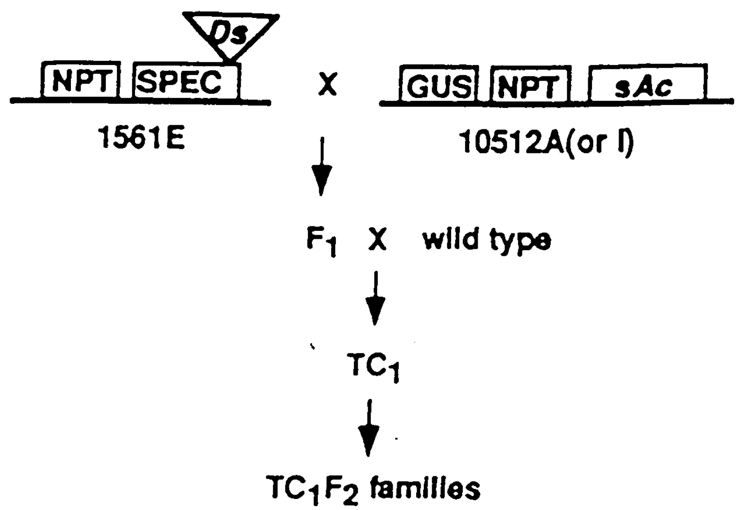
Ds
NPT
SPEC
X
GUS
NPT
sAc
1561E
10512A(or I)
F1 X wild type
TC1
TC1F2 families

FIGURE 2

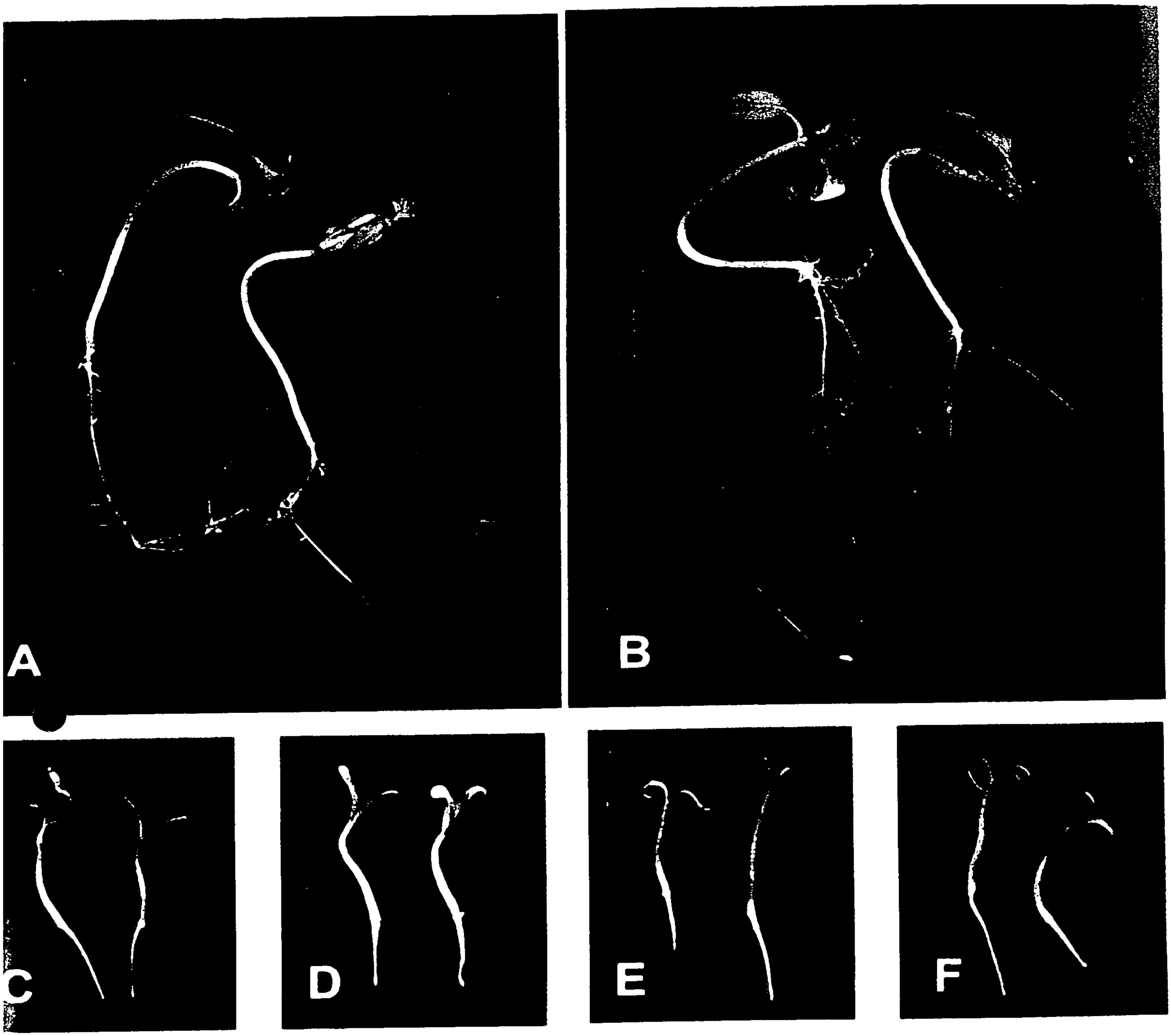

FIGURE 3

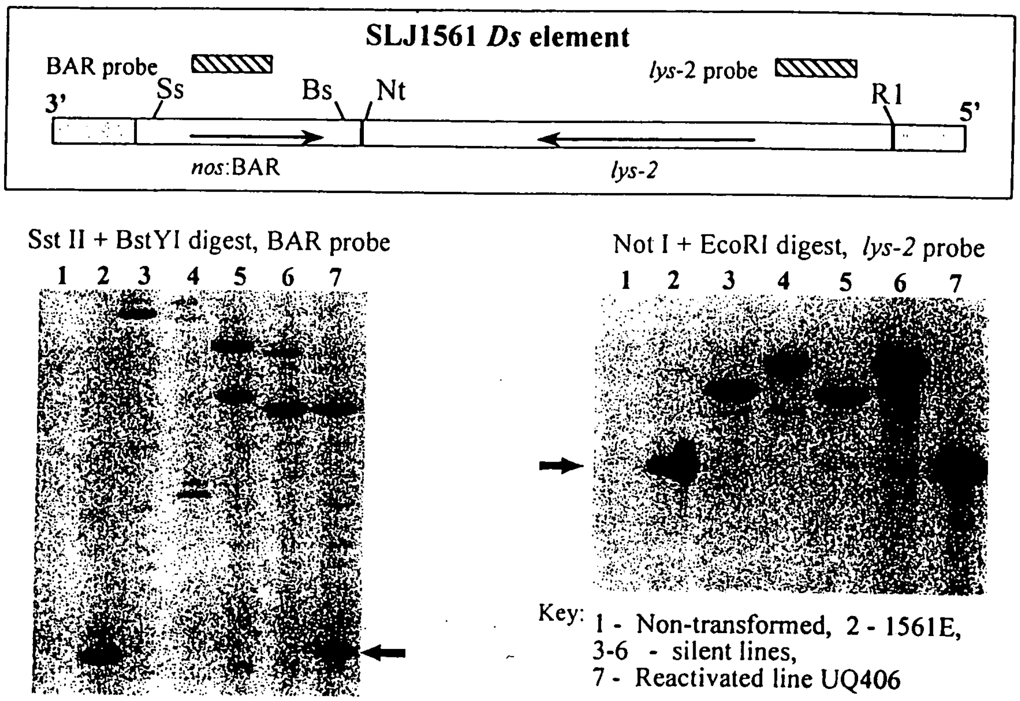
SLJ1561 *Ds* element
BAR probe
lys-2 probe
3'
Ss
Bs
Nt
R1
5'
nos:BAR
lys-2
Sst II + BstYI digest, BAR probe
1 2 3 4 5 6 7
Not I + EcoRI digest, *lys-2* probe
1 2 3 4 5 6 7
Key: 1 - Non-transformed, 2 - 1561E,
3-6 - silent lines,
7 - Reactivated line UQ406

FIGURE 4

FIGURE 5 (i)

```
 981 TTTGAAATTTATGTATATATCTGTAGCATTAGAAACTATAAGAGTTGTTA 1030  Potato
     |||||||||||||||| ||||| |||||||||||||||||||||||||||
  40 TTTGAAATTTATGTATTTATCTATAGCATTAGAAACTATAAGAGTTGTTA 89    Tomato

1031 GCTTCACTTGTCTTATTGTTGTGCTCAAAGCAACT...TCATCATACAGT 1077
     |||||||||| |||| |||||||||||||||||||   ||||||||||||
  90 GCTTCACTTGGCTTACTGTTGTGCTCAAAGCAACTTCATCATCATACAGT 139

1078 ATGGTTTTTATATGCTCTTCCATTATCACCGAACCTTATGATTATG.TGT 1126
     |||||||| |||||||||||||||||||| || ||||||||||||| | |
 140 ATGGTTTTGATATGCTCTTCCATTATCACTGAGCCTTATGATTATGTTTT 189

1127 ACGAGCTTATAATATTACTGATGGTGATTCAGTATTATGATTATGTCCTC 1176
     ||||||||||||||| |||||||||||||||||||| ||||||||||||
 190 ACGAGCTTATAATATCACTGATGGTGATTCAGTATTGTGATTATGTCCTT 239

1177 CATTAATTATTCTGTTTCATACAAGTCGTGTAATTTGCTGTTTGTGATTG 1226
     | || ||||||||||||||||||||||||||||||||||||||||||  |
 240 CGTTGATTATTCTGTTTCATACAAGTCGTGTAATTTGCTGTTTGTGACAG 289

1227 TACGATAAATTGATTCAACCTTCTGCGGTGTTGGTTGAAGTTCAAGTAAA 1276
     ||||||| || || ||||||||||| ||| || ||||||||||| |||||
 290 TACGATAGATCGACTCAACCTTCTGAGGTATTAGTTGAAGTTCATGTAAA 339

1277 TTAGCTTTATTTATCATAGTAGCATTTGATTATTGATGCTCTGTAGCTAA 1326
     |||||||| |||||||||||||||||||||||||||||||||||||||||
 340 TTAGCTTTGTTTATCATAGTAGCATTTGATTATTGATGCTCTGTAGCTAA 389

1327 TGATAAGCCATTGAAGGGAAGCAGAAATGGTAAAGCTTTCTAAAATGAAT 1376
     ||||||||||||| ||||||||          ||||||||| ||||||||
 390 TGATAAGCCATTGGAGGGAAGC..........AAGCTTTCT.AAATGAAT 428

1377 CTACGAATGGATGATAAAGTTAATGAATATTGTTGATACTTCTGCAATCA 1426
     ||||||||||||||||||||| ||||||||| ||| |||||||||| |||
 429 CTACGAATGGATGATAAAGTTCATGAATATTTTTGTTACTTCTGCAGTCA 478

1427 GATTATGAGTTACTGAGTCTACTG.TTTTTTAAGCCTGTTTCAGATGATC 1475
     ||| |||||||| |||||||| || |||||||||||||||||||||||||
 479 GATCATGAGTTATTGAGTCTATTGTTTTTTTAAGCCTGTTTCAGATGATC 528

1476 GATCATCAACAACAACATATTCAGTGTAGTAGACATGATCGATCACTTTC 1525
      |||||||  |||||||||  | |||||||   |   ||| ||||
 529 CATCATCAGTAACAACATACACGGTGTAGT..CCCAAATCCATCA..... 571

1526 TAATTTTCGATTATGCACCCTCTTTTCTCCAATTTGGTC..GTCTTCTTT 1573
                |||||||| |||||||| ||||||||||  || || |||
 572 ...........TATGCACCTTCTTTTCTTCAATTTGGTCTTGTTTTTTTT 610

1574 TTTTCATGATGTCACTGAATTATTCTCTGGTCGTCCCCACCATTCAGGAA 1623
     |||||||||||||| ||||||                    ||||| |||
 611 TTTTCATGATGTCATTGAATT....................ATTCAAGAA 640

1624 GTCACTTCGAGCATAATG...TGAAAACATCCACATTT.TTCAA...... 1663
     ||||||||||||||||||   |   || |||||| ||| |||||
 641 GTCACTTCGAGCATAATGATTTTTCAAAATCCACCTTTGTTCAAGCACTA 690
```

UQ406 insertion

FIGURE 5 (ii)

```
1664 ............ATCCAGC......................AGAATTTTC 1679
                 ||| |||                      |||||||||
 691 CCACGTCTTTTCATCTAGCCCACAACCGTGGTGGAGGATCTAGAATTTTC 740

1680 ATCAAACGGGGTTCAACATTTAC...TACATGTATACACTCTGAAGTCTG 1726
     || |||  || ||||| ||||||    | || |||||||| |  | | ||
 741 ATGAAA..GGATTCAAAATTTACAAACATATATATACACTATACACTATG 788

1727 AATCCACTAATTCTAGATGGTGCATCTGTGCCCCCACACTTGTGAAAGCT 1776
     |||||||||||| |||||||||||| |||||||||||| | ||||||||| 
 789 AATCCACTAATACTAGATGGTGCACCTGTGCCCCCACTCATGTGAAAGCC 838

1777 TATTCTCAATTTTTTATTTTCCAACAACTTGAATTCAGACCACACAACTC 1826
     |||||||||||||||||||||| ||||||| ||| |||||| ||||||||
 839 TATTCTCAATTTTTTATTTTCC.ACAACTTAAATACAGACCGCACAACTC 887

1827 CCGTGTCTTGT.......ACGGTCAGCATCTGAGTGGAGAACTCAA.... 1865
     |||||||||||       || |||||||   ||| |||||   ||
 888 CCGTGTCTTGTGTGCTCGTCGCTCAGCATGCAAGTCGAGAAAAGAAAGAC 937

1866 ..................................TTAAGTGACTTTAACG 1881
                                       || |  ||||||||||
 938 CAAAACAATGAAAACTTTACGAAAAATCAAAAAGTTGAAGGACTTTAACG 987

1882 TCGAGTTCTATAGTAAACAACCCCT..................ATATCTT 1913
     ||||| ||| | ||| | |||| ||                 |||  ||
 988 TCGAGATCTCTCGTAGAAAACCTCTTTTGTAAGGTTGCATACAATACTTT 1037

1914 TTTTCAAGCATGTTAAGATTGCGAACACACTGA................. 1946
     ||||  || |  |||    ||  |   | |||||
1038 TTTTTCAG.ACTTTACTTATGGTATTATACTGAATATGTTATTGCTGTTA 1086

1947 ........................AATTTCCAGGTCGTTAATCTTGTACC 1972
                             |||||| ||  ||||||||||||||
1087 TAGTAGTTGAGTGACGTTTGAGGGAATTTCTAGTCCGTTAATCTTGTACT 1136

1973 CAGTGTGTGTACTTTTAAAAAAAAAAGTCAGTTTTTAGTCTCTAAAACA 2022
     |||||||| |||||||      |||||||||||||||| ||||||||||||
1137 CAGTGTGTCTACTTTT...CAAAAAAGTCAGTTTTTCAGTCTCTAAAACA 1183

2023 CATTTAAAT.AGAGTTTATTTG.CCATCTTTTGTTCCTCATACTAGACTT 2070
     ||||||||| ||||||| |||| |||||||||||||||||| |||| |||
1184 CATTTAAATAAGAGTTTCTTTGCCCATCTTTTGTTCCTCATCCTAGGCTT 1233

2071 CGGAGTCAACACAACACAACAACA 2094
      |||||||||||||||||||||||
1234 .GGAGTCAACACAACACAACAACA 1256
```

FIGURE 6 (i)

1 CGACGGCCCG GGCTGGTAAA TGCGGAAGCT TGTTACAGAT TTGAAATTTA
51 TGTATTTATC TATAGCATTA GAAACTATAA GAGTTGTTAG CTTCACTTGG
101 CTTACTGTTG TGCTCAAAGC AACTTCATCA TCATACAGTA TGGTTTTGAT
151 ATGCTCTTCC ATTATCACTG AGCCTTATGA TTATGTTTTA CGAGCTTATA
201 ATATCACTGA TGGTGATTCA GTATTGTGAT TATGTCCTTC GTTGATTATT
251 CTGTTTCATA CAAGTCGTGT AATTTGCTGT TTGTGACAGT ACGATAGATC
301 GACTCAACCT TCTGAGGTAT TAGTTGAAGT TCATGTAAAT TAGCTTTGTT
351 TATCATAGTA GCATTTGATT ATTGATGCTC TGTAGCTAAT GATAAGCCAT
401 TGGAGGGAAG CAAGCTTTCT AAATGAATCT ACGAATGGAT GATAAAGTTC
451 ATGAATATTT TTGTTACTTC TGCAGTCAGA TCATGAGTTA TTGAGTCTAT
501 TGTTTTTTTA AGCCTGTTTC AGATGATCCA TCATCAGTAA CAACATACAC
551 GGTGTAGTCC CAAATCCATC ATATGCACCT TCTTTTCTTC AATTTGGTCT
601 TGTTTTTTTT TTTTCATGAT GTCATTGAAT TATTCAAGAA GTCACTTCGA UQ406
651 GCATAATGAT TTTTCAAAAT CCACCTTTGT TCAAGCACTA CCACGTCTTT insertion
701 TCATCTAGCC CACAACCGTG GTGGAGGATC TAGAATTTTC ATGAAAGGAT
751 TCAAAATTTA CAAACATATA TATACACTAT ACACTATGAA TCCACTAATA
801 CTAGATGGTG CACCTGTGCC CCCACTCATG TGAAAGCCTA TTCTCAATTT
851 TTTATTTTCC ACAACTTAAA TACAGACCGC ACAACTCCCG TGTCTTGTGT
901 GCTCGTCGCT CAGCATGCAA GTCGAGAAAA GAAAGACCAA AACAATGAAA
951 ACTTTACGAA AAATCAAAAA GTTGAAGGAC TTTAACGTCG AGATCTCTCG
1001 TAGAAAACCT CTTTTGTAAG GTTGCATACA ATACTTTTTT TTCAGACTTT
1051 ACTTATGGTA TTATACTGAA TATGTTATTG CTGTTATAGT AGTTGAGTGA
1101 CGTTTGAGGG AATTTCTAGT CCGTTAATCT TGTACTCAGT GTGTCTACTT
1151 TTCAAAAAAG TCAGTTTTTC AGTCTCTAAA ACACATTTAA ATAAGAGTTT
1201 CTTTGCCCAT CTTTTGTTCC TCATCCTAGG CTTGGAGTCA ACACAACACA
1251 ACAACAATGA ATTTCCATTT TTCTGTTTCT TTACTTCTCT CTTTATCTCT
1301 TCCTATGTTT GCCTCTTCGA CGGTGTTATT TCAGGTATCC ATCTCCAAAG
1351 AACCTTATTT TTCTCTTAAC TTTTCCTATG TATATGTATC TCTATGTTTA
1401 TGTAGTACTT GCTCAAGTAT ATAAAGAAAA GTTAGTTTCT CTAGAATCTT
1451 TGAATTCATT TGTTAGGGGT TCAATTGGGA TTCGAGTAAT AAGCAAGGCG
1501 GATGGTACAA CTCTCTCATC AACTTAGTTC CGGACTTGGC TAAAGCTGGA
1551 GTTACTCATG TTTGGTTGCC ACCATCATCT CACTCCGTTT CTCCTCAAGG
1601 TAATTTTCGG AGTGATTGTG ACCTAGTAAT CCAATGAAGT CAAAATAACC
1651 ACGGAAGATT AGAGTCTAAA TTTTAATGAA AATAGTTCAG ACAAGTTAAT
1701 GACCAACTTA TATATTAGTT CAATCCATAA AATTTGATGT AGTAGTTACA
1751 AAATGGAATT GCTTGAAGGC TTATGCCATG TTTTATGCCA GGTTATATGC
1801 CAGGAAGGTT GTATGACTAG GATGCTTCCA AGTTTGGAAA TCAGCAACAA
1851 CTGAAAACTC TTATTAAGGC TTTAACATGA CCACGGGATC AAATCGGTTG
1901 CTGATATAGT GATAAATCAT AGAACTGCTG ATAACAAAGA TAGCAGGGGA
1951 ATATACAGCA TCTTTGAAGG AGGAACATCT GATGACCGGC TTGATTGGGG
2001 TCCATCTTTC ATTTGCAGGA ACGACACACA ATATTCTGAT GGCACGGGGA
2051 ATCCAGACAC GGGTTTGGAC TTTGAACCTG CACCTGATAT CGATCATCTT
2101 AATACGAGAG TGCAGAAAGA GTTATCAGAC TGGATGAACT GGCTGAAATC
2151 TGAAATTGGA TTTGATGGTT GGCGTTTCGA TTTTGTTAGG GGATATGCAC
2201 CTTGCATTAC CAAAATTTAT ATGGGAAACA CGTCCCCGGA TTTTGCTGTT
2251 GGTGAATTGT GGAACTCTCT TGCTTATGGC CAGGACGGGA AACCGGAATA
2301 TAACCAGGAC AATCATAGAA ATGAGCTAGT TGGTTGGGTA AAAAATGCGG
2351 GGCGGGCTGT AACAGCTTTT GATTTTACAA CAAAGGGAAT TCTTCAAGCT
2401 GCAGTTCAAG AAGAGTTATG GAGATTGAAG GATCCCAATG GAAAACCTCC
2451 TGGGATGATC GGTGTTTTGC CTCGAAAAGC TGTGACTTTT ATCGATAATC
2501 ATGATACTGG ATCGACACAA AATATGTGGC CTTTCCCTTC AGACAAAGTT
2551 ATGCAAGGAT ATGCATACAT TCTTACTCAT CCAGGAATCC CATCCGTGGT
2601 AAAAAAAATA AATAAATTCT TTCTACATAT CTCATTGTTT TCTATTTTAC
2651 AAGAAATTTA TATTCTTTTC CAGGGGATTT GAGAAACTCG GCCTGTGGGA
2701 GTTTGCTCAC ATTGCCAGTC TCGTAATCCA TAAACAAACA CTCAAACTCT
2751 GAGTGTGCAC ATCTAGACAC CTCAACTCGT TTTTCACCGT GTTAATTGAA
2801 CACTTCAACT TACAAAATGA TCGTGTAGCA CCTCCAAAAA TTATGTGTCA
2851 CAATTAGCCA CGTGCGAGAT ACACGAAAAT GAGTTGGAGT AGTTAGTTGC
2901 CAAATAAAAC CAAGCTGAGG TGTCTAAATG TGCACNCTCA AAGTNGGATG
2951 TTTACTTGGC AGCTGAGGCC GAGGCCATGT TTGANTGTTA TGCTTATAGG
3001 ATATGACACA TTTGTTTCCG ATTAGCTGAG GANTTGATTA AATCCTNGTT
3051 TTNGTTNGCA GTTTNATNAC CATTNCTTTG ATNGGGGCTN CNAGGATGGA
3101 ATTNCAGCAC TAANCTCTAT TAGGAAAAGG AATAGGATTT GTGCANCAAG

FIGURE 6 (ii)

3151 CAATGTGCAA ATAATGGCTC CTGATTCTGA ATCTTTATAT ANCAATGGAT
3201 CATCACAAAA TCATTGTCAA GATTGGACCA AAACTTGATC TTGGAAATCT
3251 TATTCCACCT AATTATGAGG TGGCAACTTC TGGACAAGAC TATGCTGTAT
3301 GGGAGCAAAA GGCATAATCA TATTGTACCA CACTAAAAGG GACCATGGCC
3351 ACAATGGTTC TCATTAGTGT TAATGTTATA TGATTGAAAA TGTAATTTAT
3401 ATTGACATAA TGAAGGCCAA AAATTCAAGA AATTATAAAC AATTCAATAG
3451 TCCTTGCTCA ATTCACAATT ACATTATGAC TTCTCTATTG CAAACTAGTT
3501 TGGGTCCACA TTATTGTCTC CTAAAATTTT ACAACATTTC TTAAGGGAAC
3551 TTAATTAGTT ACAGTGAACA TATGTTGAAA TTACCCTTTA TCCCCTTACA
3601 ATTGATTTAA TAAATATTTC CCCTATCCCT TTGGTAGTTG GTTAGAGTTA
3651 TAAGTAACGT AGAGATTAGT TATAAGAGAA TTTATGTATT ATTATGCAGA
3701 TGTTTAGTTA TATCGATTTT AGTTATTTAT ATGTTGATTA TTTCACCTTC
3751 AATAATGCAT ATAAAGATGG TAAATGATTG GATTGATCGA ATTCGAATGA
3801 GTTTGAATAT GAACTAATCT TCAAATTTAA TATAAATTTT TTTTGTCAAC
3851 ATCTATAGCC AAACGGCTCC AAAACAATAA ATAATTTACA TTTATTGTAG
3901 TATTTTATTT AAAATGGGAT NTTCCTCATC CCACTTGTAC CAGTTGAAAC
3951 CCTAATAATA AGCCAATCCA ACCGTCAAAA TTACAAATTT TGAAAATTGC
4001 GCTCCTCACA GTTCTCCCCT ATTCAGATTT GATTCATTCT CTTCATTTTT
4051 TGTTTTCACA TTTTACCTCT AAATCAACAA AATTCCCTTT GTTCAAATGG Dem ATG
4101 GTGCTAATCA CAGCCGTGAA GATCTGGAGC TTTCTGATTC CGAGTCTGAA
4151 TCCGAATATG GGTCCGAGTC TCGAACAAGG GAGGAAGAGG AAGACGAAGA
4201 TAACTACTCA GATGCTAAAA CGACGCCGTC TTCCACTGAT CGGAAACAGA
4251 GCAAAACCCC GTCTTCTTTG GATGATGTTG AAGCAAAGCT GAAAGCTTTA
4301 AAGCTTAAGT ATGGTACTCC TCATGCTAAA ACCCCCACAG CGAAAAACGC
4351 TGTTAAACTT TACCTTCATG TTGGTGGGAA CACTGCGAAT TCCAAATGGG
4401 TAGTTTCTGA TAAGGTGACA GCTTATTCGT TTGTTAAATC GGGTAGTGAG
4451 GATGGATCGG ATGATGATGA AAATGAAGAA ACTGAGGAGA ATGCTTGGTG
4501 GGTTTTGAAA ATTGGGTCGA AGGTTCGGGC TAAGATTGAT GAGAATTTGC
4551 AGCTCAAGGC ATTTAAGGAG CAGAAAAGGG TGGATTTTGT GGCGAATGGG
4601 GTTTGGGCTG TGAGATTCTT TGGGGAGGAA GAGTATAAGG CGTTCATTGA
4651 CTTATATCAG AGCTGTTTGT TTGAGAATAC TTATGGGTTT GAGGCAAATG
4701 ATGAGAATAG AGTTAAGGTG TATGGTAAAG ACTTTATGGG GTGGGCAAAT
4751 CCAGAAGCTG CGGATGATTC AATGTGGGAG GATGCTGGGG ATAGCTTCGC
4801 GAAGAGCCCT GCGTCTGAAA AGAAGACACC TTTGAGGGTT AACCATGATT
4851 TGAGGGAGGA GTTTGAGGAG GCAGCTAAAG GAGGAGCTAT TCAGAGCTTG
4901 GCATTAGGTG CGTTGGATAA TAGTTTTCTT ATAAGTGATT CTGGAATTCA
4951 GGTTGTGAGG AACTATACTC ATGGAATAAG TGGAAAAGGT GTTTGTGTCA
5001 ATTTTGATAA GGAAAGGTCT GCTGTACCTA ATTCCACTCC AAGGAAAGCT
5051 CTACTTCTAA GAGCTGAGAC TAATATGCTT CTCATGAGTC CAGTGACTGA
5101 TAGAAAGCCT CACTCTCGGG GATTACATCA GTTTGATATC GAGACTGGGA
5151 AGGTTGTTAG CGAGTGGAAG TTTGAGAAAG ATGGAACTGA TATCACGATG
5201 AGGGATATCA CTAATGATAG CAAAGGAGCT CAGATGGATC CTTCGGGGTC
5251 TACTTTCTTA GGGCTAGATG ATAACAGATT GTGTAGGTGG GATATGCGTG
5301 ATCGGCATGG GATGGTCCAG AATCTAGTTG ATGAAAGTAC TCCTGTGCTG
5351 AATTGGACTC AAGGACATCA ATTTTCGAGG GGAACTAACT TTCAGTGCTT
5401 TGCTACTACT GGTGATGGAT CAATTGTTGT TGGTTCACTT GATGGCAAGA
5451 TTAGATTGTA CTCAAGCAGT TCCATGAGAC AGGCTAAAAC TGCTTTTCCA
5501 GGCCTTGGTT CTCCTATCAC TCATGTGGAT GTTACCTATG ATGGGAAGTG
5551 GATATTGGGG ACAACTGATA CTTACTTGAT ATTGATATGC ACCTTGTTTA
5601 TCGACAAGAA TGGAACTACT AAGACTGGTT TTGCTGGTCG CATGGGAAAT
5651 AAGATTTCCG CTCCAAGATT GTTAAAGCTA AACCCTCTCG ATTCACATAT
5701 GGCTGGAGCT AACAAGTTCC GCAGTGCTCA ATTTTCATGG GTCACCGAGA
5751 ATGGGAAGCA AGAGCGCCAC CTCGTTGCTA CTGTTGGGAA GTTTAGTGTG
5801 ATCTGGAATT TTCAACAGGT GAAGGATGGT TCTCATGAGT GTTACCAGAA
5851 TCAGGTTGGG TTGAAGAGCT GCTATTGTTA CAAGATAGTC CTAAGAGACG
5901 ACTCTATTGT AGAAAGTCGT TTCATGCATG ACAAGTACGC TGTTTCTGAC
5951 TCACCTGAAG CACCACTGGC GGTAGCAACC CCCATGAAAG TCAGCTCATT
6001 CAGCATCTCT AGCAGGCGCT TACAAATTTG AACAATCATT CTGTTCATAT
6051 ACGCAACTTA TTAGATTTAT CTGTAGCAGA ATTAGTGTCT CTCACACTAA

FIGURE 6 (iii)

6101 GTAGCTTGAA AAACTGCACA TCTGCAAATC ATTTCCAGTT CAATGTATTA
6151 CTACTTTAGT TTAAAAACCT TAAAAGGCAG TCTTCCAAAT TCTAGGTATC
6201 CTCACCTGAC ATTATTATTG TTGTAATAGC TAATTGTTGC TTGCTCTAAA
6251 TCCCCGTTCA ATG

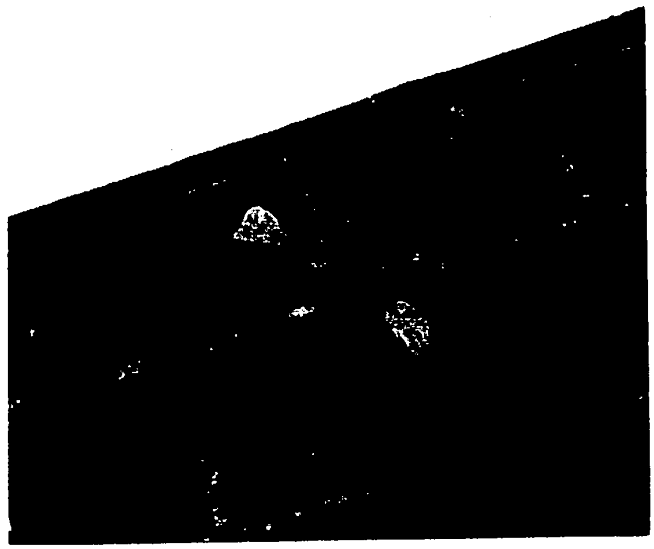

FIGURE 7

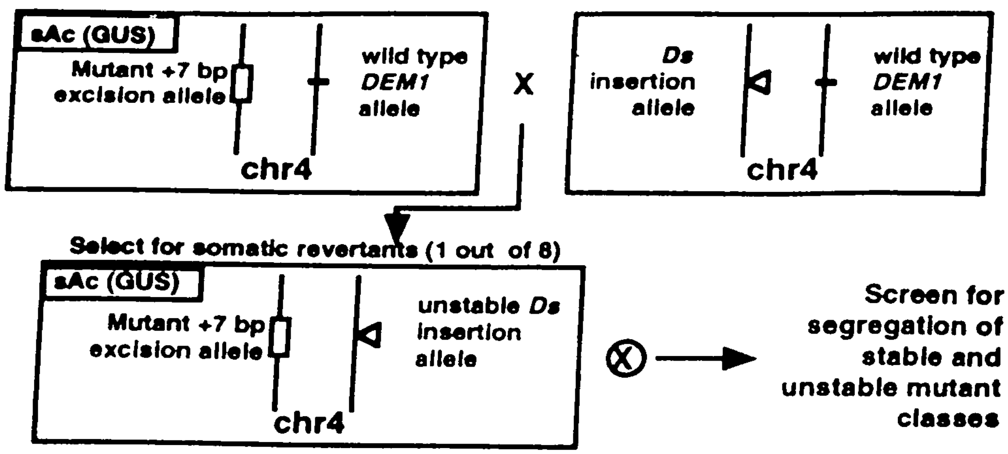
sAc (GUS)
Mutant +7 bp excision allele
wild type DEM1 allele
chr4
X
Ds insertion allele
wild type DEM1 allele
chr4
Select for somatic revertants (1 out of 8)
sAc (GUS)
Mutant +7 bp excision allele
unstable Ds insertion allele
chr4
Screen for segregation of stable and unstable mutant classes

FIGURE 8

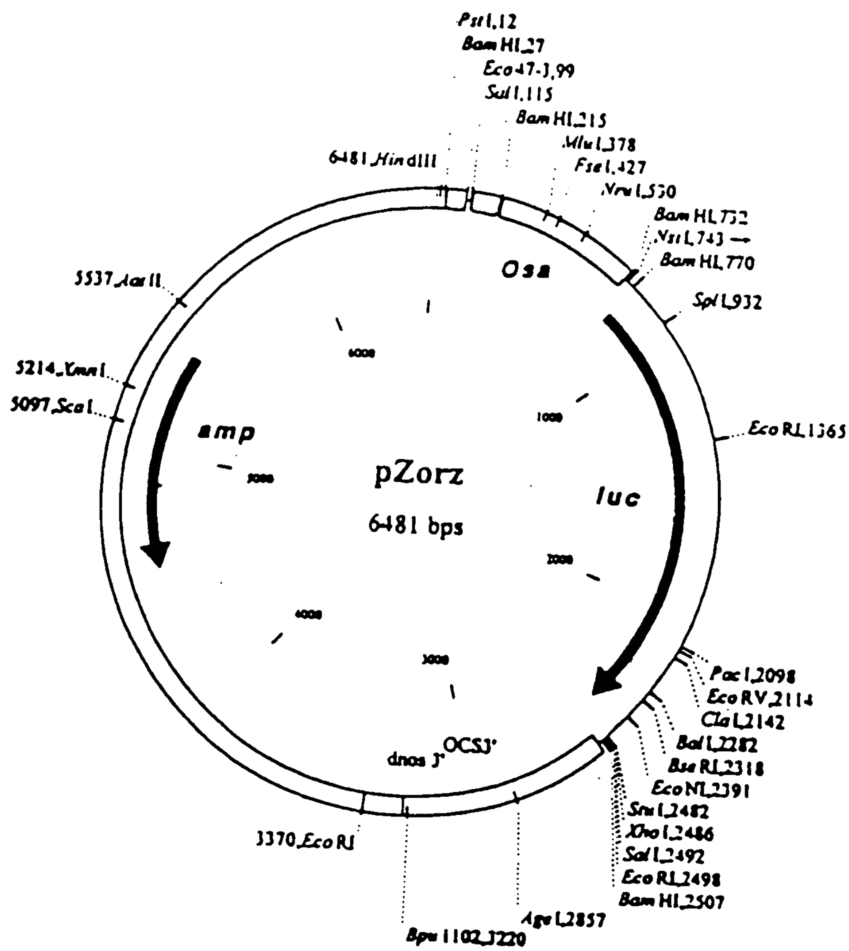
PstI,12
BamHI,27
Eco47-3,99
SalI,115
BamHI,215
MluI,378
FseI,427
NruI,530
BamHI,732
NsiI,743
BamHI,770
SplI,932
EcoRI,1365
PacI,2098
EcoRV,2114
ClaI,2142
BclI,2282
BsaRI,2318
EcoNI,2391
SnaI,2482
XhoI,2486
SalI,2492
EcoRI,2498
BamHI,2507
AgeI,2857
Bpu1102,3220
3370,EcoRI
5097,ScaI
5214,XmnI
5537,AatII
6481,HindIII
Osa
luc
amp
OCS3'
dnos 3'
pZorz
6481 bps
1000
2000
3000
4000
5000
6000

FIGURE 9